Gaming the Classroom: The Transformative Experience of Redesigning the Delivery of a Political Science Class

Mikael Hellstrom
University of New Brunswick Saint John

Abstract

Game mechanics can motivate users beyond what is normally expected. Research has shown that this technique can be used to enhance the learning experience for students on all educational levels. The paper details the experiences of transforming traditional lecture-based courses in undergraduate political science to gamification and game-based learning, and it presents the reader with a toolkit for how to make such a conversion based on the author's experiences. An overview of selected scholarly literature on teaching informs the reflection on this transformation. The paper concludes that gamification and game-based learning can provide benefits in political science education when leveraging formative assessment, flipped classrooms, and game-based learning. It also finds that there might be some institutional barriers to the adoption of these tools, primarily associated with the institutionalization of the bell curve as a guideline for the distribution of student grades. The paper ends with some reflections on possible future research areas.

Keywords: *Game-based learning; Gamification; Asessment; Political Science education*

Introduction

> What will young people come to think if they consistently see deeper learning principles in their popular culture than they do in school? (Gee, 2007, p. 218).

Much has been said about the flaws with of the traditional lecture series involving summative assessment of assignments and exams as a course design (Association for Learning Technology, 2010; Crouch & Mazur, 2001; Gibbs, 1981). Yet, this was the standard format for teaching at my department, so, like every other graduate student, I adopted it as a neophyte instructor. I quickly became frustrated, however, as I felt I had too little insight into how much my students actually learned during my courses.

During the fall of 2012, I discovered 3D Game Lab, an online teaching tool designed to facilitate gamification in education (3dgamelab, n.d.). Gamification refers to "the use of game design elements in non-game contexts" (Deterding, Khaled, Nacke, & Dixon, 2011, p. 1) and aims to increase motivation and engagement beyond what is normally expected (Fui-Hoon Nah, Rajasekhar Telaprolu, Rallapalli, & Rallapalli Venkata, 2013).

Student engagement can be defined as "student involvement or student commitment" (Hu, Ching, & Chao, 2012, p. 71) and is generally tied to positive learning outcomes. More engaged students learn more, it seems (Hattie, 2009).

When my department generously funded my attendance to the 3D Game Lab teacher camp during the winter term of 2013, I started working with it. The experience brought my attention to the literature on learning, which increasingly argues that active learning tools, student agency (i.e., when students have power over their learning process), and formative assessment produce better learning results than passive learning. However, the literature often presents the associated teaching techniques in isolation from each other.

This paper presents a reflection on lessons learned from converting a lecture-based course to a design based on gamification and game-based learning techniques. I show how instructors can merge such techniques into one comprehensive course design. The paper starts with an overview of the literature on student engagement, agency, and formative grading, then proceeds to describe how the experiences of re-designing courses can be read in light of the literature to address the questions. The course in question was Comparative Politics at the intermediate level, and this course had between 8 and 71 students in its various sections.

Notably, this paper is not describing the results of a research project comparing lecture-based courses using summative assessment with gamification and game-based learning using formative assessment. It is instead a review and interpretation of the literature based on my personal experiences. The paper ends with my wish list for such a study in light of these experiences.

Literature overview: On Student Engagement and Active Learning

The literature on learning is increasingly emphasizing student engagement as significant for learning. Student engagement occurs when students feel involved or committed to the study (Hu, Ching, & Chao, 2012, p. 71). Active learning, or learning by doing, is an important prerequisite to achieve student engagement. Using blended learning by flipping the classroom and formative assessment are some active learning methods that have been discussed recently.

Blended learning combines face-to-face learning in the classroom with digital teaching tools. The digital tools can keep track of student progress in real time, greatly facilitating the grading logistics. When properly used, students arrive better prepared to class (Bauer, 2001; Cameron, 2003), produce higher quality assignments (Benbunan-Fich & Hiltz, 1999; Garnham & Kaleta, 2002), and projects (McCray, 2000). Other positive effects have also been noted, such as the development of effective student support structures (Dziuban, Hartman, & Moskal, 2004); higher student satisfaction (Dziuban, Hartman,

Juge, Moskal, & Sorg, 2006), facilitated interaction between students and instructors (Aycock, Garnham & Kaleta, 2002), and the creation of a community of learners in the classroom (King, 2002).

'Flipping the classroom' is one way of implementing blended learning. Some describe this as the practice of recording lectures before the term starts (Ronchetti, 2010) and making them available for students online (Bergmann & Sams, 2012). This way, students can set the lecture pace individually (Owston, Garrison, & Cook, 2006), unlike in a classroom lecture where the instructor sets the pace.

Moving lectures online allows instructors to use classroom time for work. For instance, an instructor of archaeology moved the lectures online. The classroom time was instead used for an exercise in object classification (Garnham & Kaleta, 2002). Others used the freed up time for more face-to-face time with individual students (Bauer, 2001).

All of the techniques above emphasize keeping students working in the classroom. That increases student engagement (Freeman et al., 2014). The result is deeper learning (Bates & Poole, 2003; Hattie, 2009; Sorcinelli, 1991; Twigg, 2003).

Formative assessment requires students to work on a subject matter until mastery is achieved. Failure becomes a learning opportunity, as the instructor uses the early attempts as a teaching tool to show how to improve. It has been shown to be highly effective for learning, due to the high degree of feedback (Hattie, 2009). I would argue that the lowered risk associated with failure could allow instructors to use formative assessment to raise the performance expectations of students. For example, the instructor can decide that students have to display B+ or even A- level competence to get an assignment approved. The students would not be allowed to proceed to the next task until that is achieved.

In a summative assessment design, by contrast, it is more difficult to make failure a learning opportunity. Instructors might well provide students with extensive feedback after grading. However, students might disregard it. Studying feedback will have no effect on their final grades. To maximize those, they are compelled to succeed on the next exam or assignment.

Moreover, students who feel that they are in control of their learning (Alderman, 2004) are also more likely to feel engaged. That sense of control can be increased by giving students choices. An asynchronous course design, where students set their own pace through the curriculum (often a feature of online courses), provides such choice. They decide when to access learning material and complete assignments (Bates & Poole, 2003). They can access the learning tools at a time of their own choice when distractions are at a minimum and concentration can be maximized, optimizing their homework schedule. Thus, the quality of work can be increased. Instructors flipping the classroom saw positive effects from this design (Bergmann & Sams, 2012; Ronchetti, 2010), which was a key reason for student approval (Garnham & Kaleta, 2002). The flexibility also makes it easier to accommodate the needs of "individual learners" (Bates & Poole, 2003,

p. 61). For Sorcinelli, increased student choice made it became easier "to recognize the different talents and styles of learning" (1991, p. 21). Thus, performance increased (Iyengar & Lepper, 1999).

Each of the above techniques can be incorporated individually with positive effects into a course design that in all other respects is traditional. For example, a course can implement the flipped classroom but evaluate assignments using summative assessment. None of the research referenced above has made any assumptions about how the techniques might work together, nor do they reflect over how gamification or game-based learning might be relevant to this context. That is where this paper now turns to explore ways to bring the techniques together into a comprehensive design.

Converting a Lecture-based Course to Gamification and Game-Based Learning

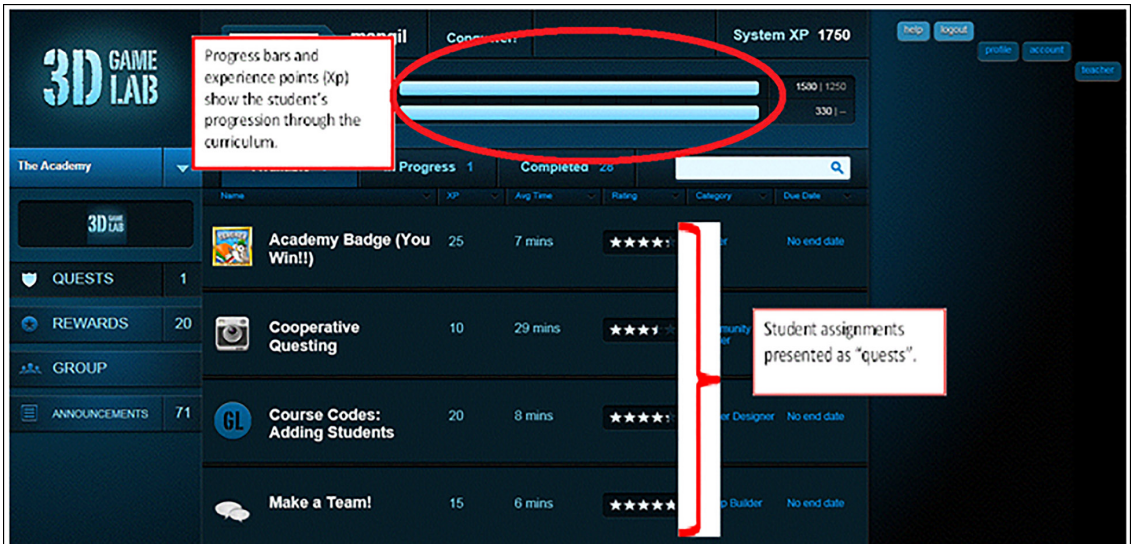
Adopting the web tool 3D Game Lab provided me with a platform where the active learning tools discussed above could be brought together comprehensively. 3D Game Lab was designed to give instructors and students in K-12 education a way to deliver gamified courses through a computer game style interface (3dgamelab, n.d.), and it has been successfully adopted by post-secondary instructors, like Davidson (2015). Gamification makes use of tools like badges, levels, and quests (Zimmerman & Cunningham, 2011). Customization includes, for example, presenting curriculum on a student's individual knowledge level or allowing students to personalize their digital user interface, and this is also common (Fui-Hoon Nah, Rajasekhar Telaprolu, Rallapalli, & Rallapalli Venkata, 2013). All these tools increase student agency and student engagement.

I converted two lecture-based courses for delivery through 3D Game Lab. The first was Political Science 230: Introduction to Comparative Politics – Global North. I taught the converted course in the Winter Term of 2013 with 71 students; in the Fall Term of 2013 with 20 students; and the Spring Term of 2014 with 21 students. The second course was Political Science 354: Topics in Comparative Politics, taught the Summer Term of 2013 with 8 students.

Turning Assignments into Quests and Experience Points into Grades

As Figure 1, below, shows, students used the interface to access and complete assignments, or 'quests'. The progress bars visualize their advancement through the curriculum, represented by the accumulation of 'experience points' (XP). The term XP is imported from digital and analogue role-playing games, where it is used as a game system for tracking a character's learning.

Figure 1. 3dgamelab interface



Using XP to represent student learning achievement can reinforce positive learning behavior. For example, Sheldon (2012) describes how he starts the syllabus presentation by telling the students that they have an F, but they can get higher grades through hard work. Students thus start the term at 0 experience points, and as they accumulate points during the term, they work themselves up through the grades. This presents the course as a challenge, motivating students to prove that they can succeed from day one. Other instructors have presented examples of variations on this theme (Glantz, 2014), and my experience of using this model was consistent with their findings.

The point accumulation system is highly compatible with formative assessment, and 3D Game Lab's computer game style user interface is designed to merge the two. This is not surprising, as many games have done so to incentivize players "…to push on through repeated failure" (Prensky, 2005, p. 113), sometimes by making failure "…interesting, and often fun" (Prensky, 2005, p. 113). Reducing the negative consequences of failure encourages the exploration of new solutions, and that facilitates learning (Gee, 2007).

This is how my syllabus explained the design to students:

> **Completing quests:** These tasks have no due date for submission, giving students maximum time to plan their own work. When a quest is completed, it will be submitted to an instructor for approval. The instructor will review the work. If the requirements have not been fulfilled, the instructor will return the quest to the student with feedback on outstanding work that needs to be completed for approval. There is no limit to the number of re-submissions a student can make. When a quest is finally approved, students will gain experience points, XP, which reflect the learning achievement (Hellstrom, n.d.)

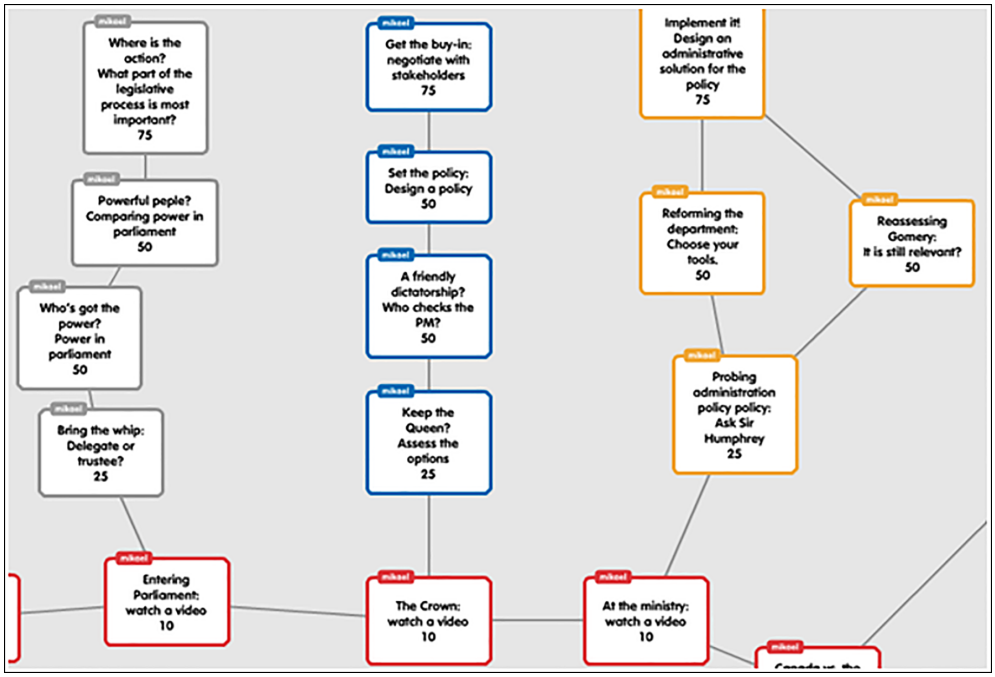
Note that the lack of due dates meant that the course was asynchronous. Students could be working on very different assignments at the same time independently of each other. That provides an opportunity for students to customize their learning schedule as

needed (Haskell, 2013b; Larsen McClarty, et al., 2012).

Any assignment the instructor considers valuable for learning can be turned into a quest. Bloom's taxonomy has shown that simpler tasks, like recalling concept names, may be insufficient for in-depth learning. More advanced tasks, such as applying concepts for analysis, or creating papers, are more conducive for deep learning (Krathwohl, 2002). I used this principle to structure my quests into pathways through the curriculum. The first assignment in such a path would often be to watch a YouTube video of a lecture introducing new concepts, for instance different types of Non-Governmental Organizations (Hellstrom, 2013). Completing that would be worth 10 XP and reveal the next task, in which the students had to use the concepts from the video to conduct an analysis. For example, the student could be instructed to find five different examples of civil society organizations on the Internet, which could be worth 50 XP. Thus, as the student ventured deeper into the topic, the assignments became more demanding.

The pathways were divided by topics, allowing students to choose the ones they were most interested in. Figure 2 shows some examples of such topics from the course 220: Introduction to Canadian Politics. The red boxes represent introductory videos, while the grey, blue, and orange ones represent in-depth specialization in the areas of parliament, executive administration, and public administration, respectively. The number in each box indicates how many experience points the student wins by completing the learning objectives for the respective quest assignment

Figure 2: Three quest-chains



In addition to quest assignments, students could also win rewards that reflect learning achievements "above and beyond the normal on the part of a student, such as in-depth

specialization in a topic, or demonstrated extraordinary abilities" (Hellstrom, n.d., p. 4). I designed these to recognize performance excellence and in-depth expertise.

Experience points and rewards provide extrinsic motivation. They are "driven mostly by the world around us, such as the desire to make money" (Zimmerman & Cunningham, 2011, p. 26). These have to be calibrated to be effective, for example "by varying the quantity and delivery schedule of that reward" (Zimmerman & Cunningham, 2011, p. 18).

Table 1, below, shows how I used rewards and experience points to calculate grades. To acquire a final grade of an 'A', a student would need to win 2000 experience points, the badge for academic writing (representing the completion of a sufficiently qualitative paper) and a second badge of the students' choice (representing the completion of all quest assignments in one course topic).

Table 1: Grade distribution schema, Introduction to Comparative Politics 230, Fall 2013

Total XP	**Letter Grade**	**Grade Point Value**	**Description**
2500 + Academic Writing Badge + any 1 badge	A+	4.0	Outstanding/ Exceptional
2000 + Academic Writing Badge + any 1 badge	A	4.0	Excellent
1750 + any 1 badge	A-	3.7	
1600	B+	3.3	Very Good
1500	B	3.0	Good
1400	B-	2.7	
1300	C+	2.3	Fully Satisfactory
1250	C	2.0	Satisfactory
1200	C-	1.7	Minimally Satisfactory
1100	D+	1.3	Adequate
1000	D	1.0	Minimally Adequate
0	F	0.0	Failure

The Logistics of Conversion

To convert the lecture-based course, I transformed my lectures to YouTube videos. I used those as basic quests to introduce students to new concepts. Exam and reader study questions were good sources of inspiration for more advanced assignments. Seventy-five quests are sufficient for a full term course, but over time, more have been added to offer students more choice.

The resulting structure replicated computer game design, where players are exposed to concepts "when players need and can use it…when the player asks for it" (Gee, 2007,

p. 218). This support model allows the student to complete a task that would otherwise be too hard. Gee's recommendations for information on-demand is also consistent with Bloom's taxonomy.

The process compelled me to shift perspective on course design. Like the instructors in Glantz' study, I started viewing the course through the "lens of the student recipient, as opposed to a previous perspective based purely on course content" (Glantz, 2014, p. 60). For example, I created a student account so I could see what students would see through the interface.

The initial redesign required a substantial effort, partly because of the need to learn the new technology. Even so, like instructors reported in the study by Aycock et al. (2002), I felt that it was well invested time, partly because I can re-use the quest database for future courses, removing the need to design new exam questions each term. That shortens the preparation time significantly. Further, as familiarity with the tools increases, the design speed increases. I now find it possible to design about 80 quest assignments and their sequence in 20 to 30 hours, excluding the production of YouTube videos. Transforming twenty lectures into videos with a simple voiceover on Prezi-presentations (Hellstrom, n.d.) took about 40 hours

Classroom Activities

The classroom was used for three types of activities: micro-lectures, prompt discussion, and quests. Micro-lectures of up to fifteen minutes addressed concepts students found particularly challenging, as identified from student submissions through 3D Game Lab. Particularly salient student questions could also prompt discussion, and a student asking such a question received an XP reward. Alternatively, the time could be used for completing quests, either individually or collaboratively. The first provided opportunity to offer nuanced feedback in person, allowing students to proceed faster through the material. The latter could be a guided lab where everyone worked on the same task.

Primarily, however, I used the classroom for game-based learning. In other words, students learned by playing games (Sheldon, 2012). When playing, learners produce content and make their own experience (Gee, 2007). As Blunt (2007, p. 4) puts it: "games require players to be part of the learning environment." Prensky argues that game-based learning can improve learning because "all games already cause players to learn" (2005, p. 105). Players learn how to do things, what to do, how to change the rules (a.k.a. 'hacking the game'), and what game rules imply about fairness and value-based decisions. They also learn about motivations, strategy and tactics, and about the game's setting (Prensky, 2005). Prensky contends that computer games are "possibly the most engaging pastime in the history of mankind" (2005, p. 102). Their feedback mechanisms encourage players to learn from past mistakes, while rewards give adrenaline and gratification. Their narratives engage player emotions and create community (Prensky, 2005). Harnessing this capacity to engage learners deeply (Larsen McClarty, et al., 2012) for educational purposes, both Blunt (2007, p. 10) and Papastergiou (2009) found that classes using games for learning resulted in higher grade point averages.

The game-based learning course design thus inherently centers on interaction. When students interact with the curriculum as players, they "draw from their own experiences and knowledge to discover facts and relationships. They interact with the world, real or imaginary, by exploring and manipulating objects and situations, wrestling with questions and challenges, and performing tasks and experiments" (Sheldon, 2012, p. 129). The design thus compels students to "construct hypotheses, make decisions, and discover principles by themselves" (Sheldon, 2012, p. 129).

While much of the literature on game-based learning focuses specifically on using computer games (Ebner & Holzinger, 2007; Gee, 2007; Papastergiou, 2009; Prensky, 2005) for learning, I did not do so in my courses. A best-selling game like Civilization V can be costly for students with tight budgets. It also has a high learning curve, which make an already substantial workload even greater.

Instead, I turned to role-playing, a form of experiential learning, which has the added advantage of providing opportunities for authentic assessment (Larsen McClarty, et al., 2012). Instructors can rarely travel to parliaments with their classes, or send students to internships in faraway political systems. Simulation brings students as close as possible to the professional environments where their skills would be used. The tool has been shown to be effective for students in higher grade levels in particular (Hattie, 2009) and used to train employees for "interviewing, communication coaching, sales, and the like" (Prensky, 2005, p. 113). The immersion and engagement provided by role-playing enhances intrinsic motivation, derived "from our core self" (Zimmerman & Cunningham, 2011, p. 26).

In political science, the Model UN constitutes a classic example of simulation, and given its popularity, it is surprising that lectures were the standard way to explore political systems at my department rather than role-plays. I developed scenarios that explored the passage of a budget through U.S. Congress, simulated a question period in the United Kingdom's House of Commons and a first minister meeting in Canada. The students submitted reports after these role-plays through 3D Game Lab, earning 75 XP for each. The reports revealed just how powerful they were for gaining an in-depth understanding of the textbook material. Students learn better when they construct ideas for themselves based on activities they have completed than when they simply listen (Prensky, 2005, p. 116). This design harnesses the notion of knowledge construction.

In summary, the gamified course design drew on many forms of active learning, including flipped learning, where students watched lectures material outside the classroom. Classroom time was used for game-based learning through role-plays. Students completed assignments through the 3D Game Lab web tool, where formative assessment of those assignments took place.

Implementing the Course Design: Reflecting on experiences in the light of active learning scholarship

I would argue that my experiences generally corroborated positive findings from previous active learning scholarship. Most students supported the gamified design, as shown in Table 2 (the results from the Summer Term 2013 course are not included as that evaluation did not contain aggregated medians). About 30% of the students made highly enthusiastic remarks in the evaluations' comments section, and only 5% were highly critical.

Table 2. Student course evaluation results, medians
Students rate instructors on a scale from 1 to 5, where 5 is the most positive

Evaluation question	Winter 2013	Fall 2013	Spring 2014
The goals and objectives of the course were clear.	4.3	4.8	4.5
In-class time was used effectively.	4.0	4.5	4.7
I am motivated to learn more about these subject areas.	4.4	4.9	4.5
I increased my knowledge of the subjects in this course.	4.6	4.9	4.7
Overall, the quality of this course was excellent.	4.3	4.7	4.8
The instructor spoke clearly.	4.7	4.8	4.7
The instructor was well prepared.	4.7	4.9	4.9
The instructor treated students with respect	4.9	4.9	4.9
The instructor provided constructive feedback throughout this course.	4.5	5.0	4.9
Overall this instructor was excellent.	4.7	4.9	4.9
The course was well organized.	4.3	4.7	4.7
The course challenged me intellectually.	4.5	4.5	4.0
The workload for this course was appropriate.	4.1	4.3	4.1
The type of assigned work was appropriate to the goals of the course.	4.4	4.9	4.5
I would recommend the course to other students.	4.5	4.8	4.9
The instructor appeared to have a thorough knowledge of the subject.	4.6	4.9	4.8
The instructor acquainted students with viewpoints other than his/her own.	4.2	4.9	4.8
The instructor assessed my work fairly.	4.4	4.9	4.8
The instructor stimulated critical thought.	4.3	4.9	4.7

Increasing Student Engagement and Choice

The asynchronous design had some interesting effects. Activity in my course dropped when other courses had midterms as students chose to study for those. I found this a net positive, as it meant that my students completed my assignments when they could give them their full attention, without distraction from other courses. Some students used this flexibility to complete the course requirements in a very concentrated fashion and accumulated 2000 XP (the threshold for an A-grade) in the first month of the course. Others achieved this in the last two weeks of the course. It would be interesting to see more research on the effects of such concentrated work on long-term retention. Students emphasized the flexible work scheduling as one of the major course design benefits. For example, one student provided this comment in the course evaluation: "The Quest-Based learning was also good, because you could go at your own pace" (Student, Political Science 230: Introduction to Comparative Politics: Global North).

To add further customization options, I developed the 'choose your medium' principle for some quests. In those, students could choose any method they wanted to communicate their understanding of the material. For example, the student could be asked to find a number of civil society organizations on the Internet and then determine whether these contributed to bonding and/or bridging social capital. Many used texts in the forms of documents or blog-posts, but some preferred visualization, like power-point presentations or Prezis, GoAnimate, or oral presentations. As long as the categorization was well motivated, the students would fulfil the objective. Thus, different learner styles were accommodated.

Moreover, students could choose which topics to pursue. The possibility of 'choosing your own path' (Student, Political Science 354: Topics in Comparative Politics), was also highly appreciated. Moreover, a few students even asked for more assignments or developed new assignments themselves (and were rewarded accordingly with XP), which they shared with peers. This is arguably a manifestation of the sense of accomplishment called the IKEA-effect, where even an arduous solitary task can "induce greater liking for the fruits of one's labor" (Mochon & Norton 2012, p. 453). This might be an effect of students engaging more deeply with the material on the higher cognitive levels of Bloom's Taxonomy, which are analysis and artefact creation. All of these customization options provided students with more choice, and their comments echo Gee's words about how game-design in learning helps students "feel a real sense of agency, ownership and control" (Gee, 2007, p. 217).

The classroom became a network with a research capacity surpassing any one individual, and students routinely found resources that would otherwise have gone unnoticed. I felt that the level of sophistication in classroom conversations was higher than I had experienced in my older course designs. At least some students shared this feeling: "I enjoyed the seminars greatly and found I learned a lot from them. The quest based learning was the most, though. Created an excellent dialogue and enhanced my understanding" (Student, Political Science 230: Introduction to Comparative Politics: Global North).

The customization options were particularly helpful for students with educational barriers. International students struggling with English found this format less intimidating: "As an international student, sometimes I feel very nervous because [the professor] required a lot of participation in class discussion, but finally, I could find myself grow and learned a lot from this course" (Student Political Science 230: Introduction to Comparative Politics: Global North). Students with disabilities and mental health issues welcomed it for similar reasons. For example, I had students with social anxiety who were very pleased with the possibility to complete assignments from home when the classroom setting became too intimidating.

However, I cannot comment on the effects across some demographic categories like race, gender, or age. For the last group, the potential effects of a 'digital generational divide' between young and seniors may well have particular significance for this design's potential, but I have had too few students approaching senior age to comment on whether such students would find this design particularly challenging.

Improving the learning experience

Students provided constructive critique on how to improve the design. Mostly, they asked for more clear assignment objectives and explanations for how assignments were connected, as well as a decrease in workload. As Table 2 shows, the workload did get the lowest grade in the student evaluation. One student wrote: "…even 400-level courses I took didn't ask for this much work" (Student, Political Science 230: Introduction to Comparative Politics: Global North). The workload was, indeed, higher, mostly because of the formative assessment (see further below). Still, students gave an assessment of above 4.0 on the course evaluations, so there was some tolerance, possibly an effect of higher levels of engagement.

Further, the design seems to have increased the transparency of student performance assessment. For example, 3D Game Lab's progress bar, previously shown in Figure 1 above, might look like a gimmick, but it is both an indicator of how well they have performed in the past and also how much work they need to complete to qualify for higher grades. As one student put it, it "helped keep track of what I learnt and what I still needed to do as well as a good way to keep track of my grade, makes the student responsible for learning" (Student, Political Science 230: Introduction to Comparative Politics: Global North). Such transparency is important because students can "shape their work intelligently and appropriately while it is being developed" (Sadler, 2005, p. 178).

In addition, 3D Game Lab provides instructors with an automatically updated spreadsheet, which gives immediate information about what students were working on and what they had completed. Students could also rate assignments, providing information about assignment design quality. I used this data to calibrate assignments and classroom presentations.

Significantly, formative assessment helped me identify students who struggled with writing, so I could give them more of my attention. To illustrate, during the Fall Term of 2013, I had a student whose first attempt at writing an introduction to a paper was clearly substandard. With summative assessment, the work would have received an F or a D. In this design, I returned the paper with comments on how to improve it. It took about five drafts to improve the paper. The final product was a B, maybe even B+, level. Other students needed up to eight drafts. My impression is that this is the most powerful tool I have used to date to support students struggling with writing.

Moreover, this use of formative assessment mimics how academic writing works on higher levels. The political science graduate thesis writing process involves having the student submit drafts to the supervisor, who provides feedback. The work continues in an iterative fashion until the thesis has sufficient quality for defense. The peer review works according to similar principles, at least in theory. Formative assessment is thus a reflection of authentic academic work. I would therefore argue that it should be explored as a pedagogical tool for teaching at the undergraduate level in political science.

Similar improvements were noticeable in overall course performance. Several students displayed an initial performance that would have resulted Ds or Cs in my summative assessment course designs, and this would have left them without possibility to turn their failure into a learning opportunity. In the 3D Game Lab course design, students did have that opportunity, and they managed to achieve grades ranging from C to A+. Some even managed to compensate for poor performance earlier in the term by working hard toward the end of the term, going from a grade of F two weeks before the deadline to, in a couple of cases, A-level grades. This experience has been consistent throughout the courses I have taught.

The research on learning often emphasizes the significance of extensive and timely feedback for learning. Improvements here might well have been a result of such feedback, though a comprehensive study is needed to explore this further by checking students' writing ability at the start of the term, to track learning trends over time for specific writing assignments, and to use control groups to isolate variables. Students emphasized that the design gave them unprecedented feedback levels both in terms of how fast it was provided, often within twenty-four hours of completing an assignment, and the level of detail, and that this was helpful for them. This is a typical reaction: "I loved the immediate feedback" (Student, Political Science 230: Introduction to Comparative Politics: Global North). These conversations were useful for me as well, providing me with important insights about student skill levels. Thus, I could identify what weaknesses needed addressing for each student.

Increasing the volume of communication between instructor and individual students raises the intuitive concern that this design is more time consuming than summative assessment designs. While that concern is understandable, there are reasons not to jump to conclusions. First, the most time consuming conversations are with the students who struggle the most with the material, and these are the students with whom instructors

should be spending the most time. Second, many assignments ask the student to produce a couple of paragraphs of stream of consciousness text, which only takes a few minutes to peruse. Third, in my summative assessment designs, my teaching team often needed one or two weeks when grading midterms or papers without being able to address learning gaps. Further, in the asynchronous design, assignments tend to come in more continuously during the term, so time devoted to assessment is less concentrated. Towards the end of the term, the workload increased somewhat as a significant minority of the students in each class tended to wait until quite late with coursework, which affected turn-around time negatively. How the design affects time use in comparison to approaches that are more traditional may need to be established through a comparative research project.

Noticed effects on grades, grading practices and time use

Outcomes suggest that the time invested had positive effects on learning achievements. One of 3D Game Lab's developers reported that in his post-secondary education class:

> …93% of students (N=97) …reached the winning condition, described as receiving a course grade of 'A'...In this approach, average completion time was reduced from 16 weeks to 12 ½ weeks with one student completing in just four weeks and many of the students who reached an A-grade…continued playing through the curriculum, demonstrating persistence in learning…Not only did the vast majority of students reach the winning condition, many exceeded expectations. As a group, the class averaged nearly twice as many completed activities as previous, module-based iterations of the course (Haskell, 2013b).

The students even kept working well beyond the point at which they achieved an A because they wanted to accumulate more points (see Figure 3). That suggests a very high level of motivation and engagement among students. Likewise, Davidson's use of 3D Game Lab resulted in 100% A's (2015).

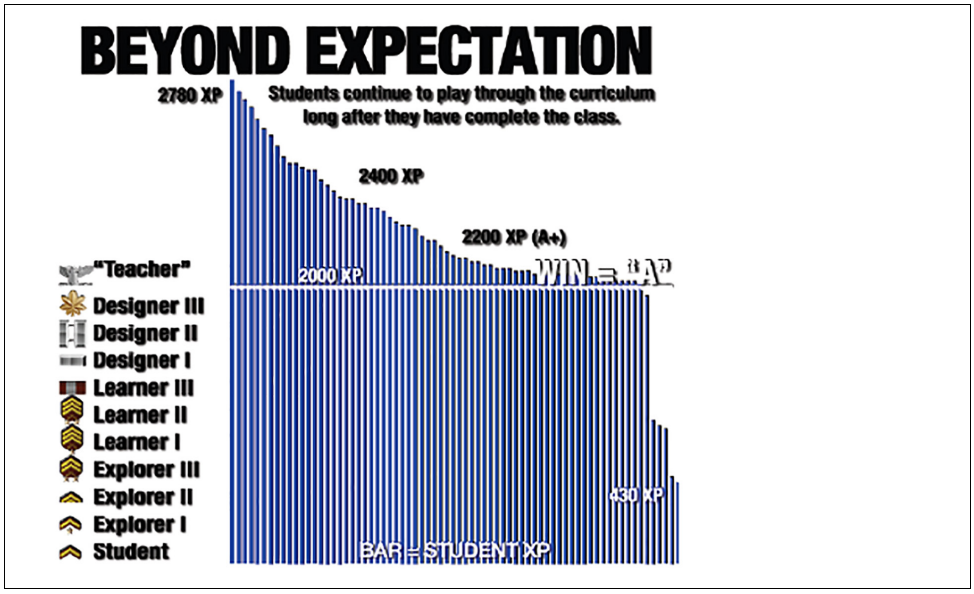


Figure 3.
Haskell C., 2013. Understanding Quest-Based Learning: Creating effective classroom experiences through game-based mechanic and community. Whitepaper. Figure 4: "Pre-service teacher candidates level up and remain persistent after earning 'A'", p. 4.

My students did not perform at that level, but improvements were still made compared to previous cohorts. Table 3 below shows the grade average results. For Political Science 354: Topics in Comparative Politics, there was no discernible difference between student performances before or after the design change. For Political Science 230: Introduction to Comparative Politics – Global North, the average grade reached almost a full score higher than previously recorded.

Table 3.
Class Grade Averages before and after the redesign

Courses	**Before the redesign Grade Average**	**After the redesign Grade Average**
Political Science 354: Topics in Comparative Politics	B+	B+
Political Science 230: Introduction to Comparative Politics – Global North	B	Winter 2013 B+ Fall 2013 A- Spring 2014 B

T-tests reveal that there is no statistical difference between the classes of Fall 2012 (28 students) and Winter of 2013 (71 students) with a p-value of $p=0.19$. Comparing Fall 2012 with Fall 2013 (20 students) did reveal a statistical difference, with $p=0.001$. Future research could shed more light on these differences.

Some have reservations about high grade averages and argue that the A-level grades should be reserved for the particularly talented, but, then, what constitutes talent? Instructors assign grades based on performance, and I wonder if A grades tend to be assigned to students with more experience with the field, rather than the talented. At least, that was what happened in my previous designs; A-level students were often third or fourth year students taking 100 and 200-level courses as part of their degree requirement and had more experience with social sciences than their junior colleagues had. A formative assessment format allows junior students an otherwise unavailable chance to develop experience.

After the Fall 2013 course, the department grew concerned that the improvements were a result of grade inflation and told me that class grade average had to be lowered. I requested a review of the students' work, to ascertain whether my assessment was lacking sufficient rigor to differentiate among students. Davidson encountered a similar reaction to her grade outcomes and referred to it as the 'Battle of the Curve'. In her case, the department reviewed the student artifacts and found that the grades were, in fact, a proper reflection of student performance (2015). The department has since supported her successful application for research grants to explore gamification for teaching further (Centre for Teaching and Learning, 2014).

That review did not happen in my case. To satisfy the department, I set the mastery level at a B-grade, with A-level grades reserved for students with awards for excellence. The reduction in grade average for the Spring Term course of 2014 was the result. There were two problems with this solution. First, students who knew they could not improve their grades lost motivation and thus stopped learning. Second, my impression, though this needs verification through further research, is that students who wanted high grades responded to the new demands by working harder. While this was a successful strategy for some, one has to ask at what point the workload becomes unreasonably high even for A-level grades.

Grade distribution also changed. Instead of following the traditional bell curve shape, it clustered around grade thresholds, as students who did not meet the requirements for achieving a certain grade 'got stuck' just below it, while others could move well past that grade, as shown in Table 4. One important question that warrants further discussion concerns what would constitute mastery in political science.

Table 4.
Course Grade Distribution – percentage of students achieving each grade level.

Grade Level	Winter 2013	Summer 2013	Fall 2013	Spring 2014
A+	13%	25%	30%	14%
A	25%	13%	15%	5%
A-	7%	0	35%	14%
B+	31%	25%	10%	38%
B	7%	25%	5%	10%
B-	4%	0	0	5%
C+	1%	0	5%	0
C	0	0	0	0
C-	0	13%	0	0
D+	0	0	0	5%
D	4%	0	0	0
F	7%	0	0	14%

A Wish List for Future Research on Gaming the Political Science Classroom

Gamification and game-based learning have been explored across educational levels and subjects. At the post-secondary level, their effects on nursing (Davidson, 2015), computer science (Sheldon, 2012), education (Haskell, 2013b), and civil engineering (Ebner & Holzinger, 2007) have been discussed. My experiences of using 3D Game Lab in political science have given me reason to believe that it holds great potential in this field as well. Even so, the literature discussed above has yet to answer a series of questions.

Many students said that the course design was refreshing and that others should adopt it. How much of that engagement was driven by novelty? I would not be surprised if novelty plays some role, but I could not speak to how large it would be. Well-designed computer games can retain player interest over long periods. That observation encouraged scholars like Gee, Haskell, and Sheldon to pursue research on this type of pedagogy (Gee, 2007; Haskell, 2013b; Sheldon, 2012).

How does the engagement effect vary across student demographics? I have already commented on how students with educational barriers found it particularly helpful. That said, it is entirely possible that some students benefit more than others across demographic variables, like age, gender, or socioeconomic status. There is a possibility that students who are also gamers benefit more than those who are not gamers.

As mentioned, some students managed to accumulate 2000 XP in a matter of weeks. Did they gain the same level of in-depth knowledge as those who took longer to meet the requirements? The classic story of the student who studies hard only during the final days before the exam, achieves a good mark, but then forgets most of the material a few weeks later comes to mind, gives reason to ask this question. A research project would have to maintain contact with students and explore knowledge retention several months after the course ended to answer this. The project should also compare performances to students in a lecture-based course relying on summative assessment.

Class size is another important dimension. The largest class covered here had 71 students. What about classes larger than 100 students? Previous experience has shown that it is hard to implement formative assessment in large classes (Owston, Garrison, & Cook, 2006). Comparing how much time this design consumes relative to a traditional lecture-based course with summative assessment would be useful.

Another issue concerns the course design's compatibility with institutionalized practices, which may act as barriers for educational innovation regardless of pedagogical potential. Many faculty members care deeply about quality teaching, but they live under time constraints, without the resources for implementing new teaching tools. This has been discussed for decades (Gibbs, 1981).

Normative assessment is another problematic feature. It is based on the principle of sorting "diamonds from the dirt" (Haskell, 2013a), where students are evaluated in comparison to each other. As a result, the number of grades of each level effectively becomes rationed, with only so many A grades 'up for grabs'. The model has been said to protect "the values of grades throughout the institution and over time… reducing any tendency towards grade inflation" (Sadler, 2005, p. 187). The above has shown the grade results in the classes where 3D Game Lab, with more than 90% of Davidson's and Haskell's students receiving an A-level grade (Davidson, 2015; Haskell, 2013b). Such high averages contradict the principles of norm-based assessment, and my department reacted negatively to the outcomes. It is possible that other departments would react similarly.

Still, there is an opportunity here. Norm-based assessment has been found to be flawed. In practice, it only reflects who won 'the race'. It does not reflect how much the students learned, nor does it reveal "poor course design, poor teaching or poor assessment processes or tasks. Conversely, excellence in course design, teaching and assessment equally go unrecognised" (Sadler, 2005, p. 187). As departments increasingly move away from grading on a curve, the opinion that students "deserve to be graded on the basis of the quality of their work alone" (Sadler, 2005, p. 178) becomes more common. A conversation on these practices could also reflect on the potential of formative assessment.

One such conversation could address what should be considered mastery in political science. Should it be an A-level grade or a B-level grade? How does the determination of that threshold affect student motivation? Is it a good idea to have a threshold that might discourage students from further work once it has been achieved? These are some of the questions that might be asked in future studies

Conclusion

In summary, pedagogical research on active learning has shown positive effects on learning. For example, when Crouch and Mazur tracked learning outcomes from peer learning, they found that student scores "improved dramatically" compared to the lecture series design previously used (Crouch & Mazur, 2001, p. 975). However, each of the active learning methods has been investigated in isolation from the other. What is thus the potential benefit of bringing these tools together? The development of a comprehensive set of metrics of the same validity and reliability produced by Crouch & Mazur (2001) would provide a good point of departure for further discussion on the veracity of gamification and game-based learning as a platform for leveraging many active learning tools in a single course design.

References

3dgamelab. (n.d.). Retrieved September 23, 2014, from 3dgamelab: http://3dgamelab.com/

Alderman, M. K. (2004). *Motivation for achievement: possibilities for teaching and learning* (2nd ed.). Mahwah: Lawrence Erlbaum Associates.

Association for Learning Technology. (2010, September 29). *Don't lecture me - Donald Clark at ALT-C 2010*. Retrieved May 21, 2015, from Youtube: https://www.youtube.com/watch?v=Tbl-xXF8NPY

Aycock, A., Garnham, C., & Kaleta, R. (2002). Lessons learned from the hybrid course project. *Teaching with Technology Today*, 8(6), 9-21.

Bates, T., & Poole, G. (2003). *Effective teaching with technology in higher education: Foundations for success*. San Franscisco: Jossey-Bass.

Bauer, W. (2001). Enriching the traditional music classroom through Internet-based technologies. *The Technology Source* (July/August). Retrieved July 3, 2014, from http://technologysource.org/article/enriching_the_traditional_music_classroom_through_internetbased_technologies/

Benbunan-Fich, R., & Hiltz, S. R. (1999, March). Educational Applications of CMCS: Solving Case Studies through asynchronous learning networks. *Journal of Computer-Mediated Communication*, 4(3). Retrieved July 3, 2014, from http://onlinelibrary.wiley.com.login.ezproxy.library.ualberta.ca/doi/10.1111/j.1083-6101.1999.tb00098.x/full

Bergmann, J., & Sams, A. (2012). *Flip your Classroom: Reach Every Student in Every Class Every Day*. Eugene: ISTE.

Blunt, R. (2007). Does Game-Based Learning Work? Results from three recent studies. *The Interservice/Industry Training, Simulation & Education Conference* (I/ITSEC). National Training Systems Association.

Cameron, B. (2003, September/October). The effectiveness of simulation in a hybrid and online networking course. *Techtrends*, 47(5), 18-21.

Crouch, C. H., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics*, 69(9).

Davidson, S. (2015). Changing "The Game" in Higher Education: Responding to Resistance. QuestBoise 2015: *Quest-Based Learning Virtual Unconference.* Boise State University. Retrieved March 30, 2016, from http://www.questboise.com/#!2015-speakers/ch2k

Deterding, S., Khaled, R., Nacke, L. E., & Dixon, D. (2011). Gamification: Toward a definition. *CHI 2011 Gamification Workshop Proceedings*. Vancouver.

Dziuban, C. D., Hartman, J. L., & Moskal, P. D. (2004). Blended Learning. *Educause Center for Applied Research* (7). Retrieved March 30, 2016, from https://net.educause.edu/ir/library/pdf/ERB0407.pdf

Dziuban, C., Hartman, J., Juge, F., Moskal, P., & Sorg, S. (2006). Blended learning: Online learning enters the mainstream. In C. J. Bonk, & C. R. Graham (Eds.), *The handbook of blended learning: global perspectives, local designs.* San Franscisco: Jossey-Bass.

Ebner, M., & Holzinger, A. (2007). Successful implementation of user-centred game based learning in higher education: An example from civil engineering. *Computers & Education*, 49, 873-890.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23), 8410-8415.

Fui-Hoon Nah, F., Rajasekhar Telaprolu, V., Rallapalli, S., & Rallapalli Venkata, P. (2013). Gamification in Education Using Computer Games. In S. Yamamoto (Ed.), *Human interface and the management of information. information and interaction for learning, culture, collaboration and business: 15th international conference, HCI international. 8018*, pp. 99-107. Berlin Heidelberg: Springer.

Gamson, Z. F. (1991). A Brief History of the Seven Principles for Good Practice in Undergraduate Education. In A. W. Chickering, & Z. F. Gamson (Eds.), *Applying the seven principles for good practice in undergraduate education: new directions for teaching and learning* (Vol. 47). San Francisco: JosseyBass.

Garnham, C., & Kaleta, R. (2002). Introduction to Hybrid Courses. *Teaching with Technology Today*, 8(6). Retrieved March 30, 2016, from http://www.uwsa.edu/ttt/articles/garnham.htm

Gee, J. P. (2007). *What video games have to teach us about learning and literacy* (2nd ed.). New York: Palgrave: Macmillan.

Gibbs, G. (1981). Twenty terrible reasons for lecturing. *SCED Occasional Paper* No. 8, Birmingham. Retrieved September 26, 2014, from http://www.brookes.ac.uk/services/ocsld/resources/20reasons.html

Glantz, E. J. (2014). Higher Education Classroom Community Game. In *Innovative Practices in Teaching Information Sciences and Technology* (pp. 53-62). Springer International Publishing.

Haskell, C. (2013a). Blowing up the grade book. *Quest Boise 2014 Quest-Based Learning Virtual Unconference*. Boise: Boise State University. Retrieved May 11, 2015, from http://www.questboise.com/#!2013-videos/c1gz6

Haskell, C. (2013b). *Understanding Quest-Based Learning: Creating effective classroom experiences through game-based mechanics and community.* Department of Education. Boise State University.

Hattie, J. (2009). *Visible Learning: A synthesis of over 800 meta-analyses relating to achievement.* New York: Routledge.

Hellstrom, M., [Professor Hellstrom] (2013, May 2). *Civil society scholarship- an overview*. Retrieved May 14, 2015, from YouTube: https://www.youtube.com/watch?v=AwP4BOLIuRU

Hellstrom, M., (n.d.). *Professor Hellstrom*. Retrieved December 21, 2016, from https://www.youtube.com/user/MrMHellstrom

Hellstrom, M., (n.d.). Syllabus, Political Science 463: Topics in Post-Industrial States, Winter 2015. University of Calgary, Department of Political Science.

Hu, Y. L., Ching, G. S., Chao, P.C. (2012). Taiwan student engagement model: Conceptual framework and overview of psychometric properties. *International Journal of Research Studies in Education*, 1(1), 69-90.

King, K. P. (2002). Identifying success in online teacher education and professional development. *The Internet and Higher Education*, 5(3), 231-26.

Krathwohl, D. R. (2002). A revision to Bloom's Taxonomy: An overview. *Theory into Practice*, 41(4).

Larsen McClarty, K., Orr, A., Frey, P. M., Dolan, R. P., Vassileva, V., & McVay, A. (2012). A literature review of gaming in education. *Gaming in Education*, 1-35.

McCray, G. E. (2000). The hybrid course: Merging on-line instruction and the traditional classroom. *Information Technology and Management*, 1(4), 307-327.

Mochon, D., & Norton, M. I. (2012). The IKEA effect: When labor leads to love. *Journal of Consumer Psychology*, 22(3), 453-460.

Owston, R. D., Garrison, D. R., & Cook, K. (2006). Blended learning at Canadian universities. In C. J. Bonk, & C. R. Graham (Eds.), *The handbook of blended learning: Global perspectives, local designs* (pp. 338-350). San Francisco: Pfeiffer.

Papastergiou, M. (2009). Digital game-based learning in high school computer science education: Impact on educational effectiveness and student motivation. *Computers and Education*, 1-12.

Prensky, M. (2005). Computer games and learning: Digital game-based learning. *Handbook of computer game studies*, 18, 97-122.

Ronchetti, R. (2010). Using video lectures to make teaching more interactive. *International Journal of Emerging Technologies in Learning*, 5(2), 45-48.

Sadler, R. D. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment and evaluation in higher education*, 175-194.

Sheldon, L. (2012). *The multiplayer classroom: Designing coursework as a game*. Boston: Cengage Learning.

Sorcinelli, M. D. (1991). Research Findings on the Seven Principles. In A. W. Chickering, & Z. F. Gamson (Eds.), *Applying the seven principles for good practice in undergraduate education: New directions for teaching and learning* (pp. 13-26). San Francisco: JourneyBass.

Centre for Teaching and Learning. (2014). *TLEF 2014-2014 adjudication results*. Retrieved May 8, 2015, from: http://uofa.ualberta.ca/centre-for-teaching-and-learning/tlef/tlef-projects/tlef-2014-15-adjudication-results

Twigg, C. A. (2003). Improving quality and reducing cost: designs for effective learning. *Change*, 35(4), 22-29.

Zimmerman, G., & Cunningham, C. (2011). *Gamification by design: implementing game mechanics in web and mobile apps*. Sebastopol: O'Reilly Media